VIROLOGY 245, 128–141 (1998)
ARTICLE NO. VY989136

Identification of Transcripts and Promoter Regions of Ovine Adenovirus OAV287

Aparajita Khatri and Gerald W. Both[1]

Molecular Science, CSIRO, North Ryde, New South Wales 2113, Australia

Received December 29, 1997; returned to author for revision February 10, 1998; accepted March 9, 1998

The ovine adenovirus isolate OAV287 represents a new group of adenoviruses that are distinct from the Mast- and Aviadenoviruses by several criteria, including genome arrangement. The OAV major late promoter and some late transcripts were previously mapped. To better define the probable coding sequences and to identify the approximate location of early promoters a partial transcription map of the genome was elucidated using a PCR-based approach. This was possible because the complete nucleotide sequence of the genome was known. The strategy permitted the identification of transcription start sites and RNA splice junctions and allowed the approximate location of promoters in the lefthand end, IVa_2, E2, P32K, and E4 regions to be deduced. The data showed that lefthand end and E4 regions are controlled by three and two temporally distinct promoters, respectively. The E2 region is controlled by a single promoter, in contrast to Mastadenoviruses, where E2 expression is controlled by the E2A and E2B promoters. The p32kDa structural protein at the lefthand end and the IVa_2 protein are also expressed from their own promoters. These data contribute to the first overview of transcription from a non-*Mastadenovirus* genome.

INTRODUCTION

The Adenoviridae have previously been considered to consist of two genuses, *Mastadenovirus* (including all known human and most mammalian viruses) and *Aviadenovirus* (typified by CELO virus) (Norrby *et al.,* 1976). More recently, a third group of adenoviruses (Ad) (as yet unnamed) was identified by phylogenetic analysis of viral protease genes. This group includes bovine Ad7, egg drop syndrome (EDS), and an ovine Ad (OAV287) (Harrach *et al.,* 1997). The OAV (Vrati *et al.,* 1995, 1996a, b) and EDS genomes (Hess *et al.,* 1997) have now been sequenced and compared with other viruses, including Ad2 (Roberts *et al.,* 1984) and CELO virus (Chiocca *et al.,* 1996), the prototype Mast- and Aviadenoviruses, respectively.

The genome of Ad2 is divided into five early transcription units, EIA/B, E3, E4, driven by individual promoters, E2, controlled by the E2A and E2B promoters, and one late transcription unit that is controlled by the major late promoter (MLP). From each promoter, RNA polymerase II generates families of transcripts which are modified by extensive splicing. In addition, there are two delayed early transcription units, IVa_2 and IX, that are also controlled by their own promoters. The E1, IX, MLP, and E3 units are transcribed from left to right on the top strand of the genome, while the E4, E2, and IVa_2 units are transcribed from right to left on the opposite strand, in that order, respectively. Among the Mastadenoviruses, one or two VA RNA genes may also be located on the top strand between the 52/55K and terminal protein genes (reviewed in (Horwitz, 1990; Shenk, 1996).

Comparison of the genomes of OAV (Vrati *et al.,* 1995, 1996a, b), CELO (Chiocca *et al.,* 1996), and Ad2 (Roberts *et al.,* 1984) reveals that each virus group has a genome structure that is distinct and characteristic. The central core of each genome, flanked by the IVa_2 and fiber genes, contains sequences that encode the majority of structural proteins and the DNA replication proteins, but the left- and righthand ends (LHE, RHE) have open reading frames that are unique to each group. OAV lacks a typical E1A/E1B region (as defined by the Mastadenoviruses) at the LHE, and has a putative E4 region that is penultimate in the genome to a group of reading frames of unknown function. These are located at the RHE on the bottom strand. Both the E4 and RHE regions appear to be independent transcription units (Vrati *et al.,* 1996b; Xu *et al.,* 1997). In addition, OAV lacks sequences coding for protein IX and core protein V. However, a new structural protein that appears to require its own promoter is encoded on the bottom strand at the LHE of the genome (Vrati *et al.,* 1996b). Consequently, the transcription maps established for the Mastadenoviruses (reviewed in (Horwitz, 1990) will differ from the other two groups, particularly for the regions encoding nonstructural proteins. We therefore initiated this work to gain an overview of transcription from the OAV genome. This is the first such work for a non-*Mastadenovirus* genome.

Because the OAV genome sequence was known, our strategy to elucidate a transcription map was to reverse

[1] To whom correspondence and reprint requests should be addressed at CSIRO, Molecular Science, PO Box 184, North Ryde, NSW, 2113, Australia. Fax: 61 2 9490 5005. E-mail: gerry.both@molsci.csiro.au.

0042-6822/98 $25.00

transcribe RNAs from OAV-infected cells and use primers to amplify them by PCR. cDNAs were cloned and sequenced to identify exons and introns in spliced transcripts and to define probable coding sequences. Additional analyses were carried out to identify transcription start sites so that the probable location of promoters in the OAV genome could be inferred. Temporal expression from some promoters was also analysed. RNAs from all transcription units, as well as seven new promoter regions, were identified.

RESULTS

As a general strategy to identify transcripts, primers that fell within an open reading frame (ORF) were used in conjunction with upstream primers that were located near candidate promoters. Alternatively, the ORF primer was used together with the 5′ anchor primer in RACE–PCR to amplify sequences that included the 5′ end of the RNA. cDNAs were cloned and sequenced to confirm their identity and to identify exons and introns.

Transcription from the LHE

The revised sequence of the OAV genome (GenBank Accession Nos. U18755, U31557, U40839, and U40837) predicts that the first 3 kb at the LHE contains three main reading frames on the top strand (ORFs LH1, 2, and 3) which potentially code for proteins of 9.7, 14.7, and 42.8 kDa (Fig. 1A and Table 1), respectively. The bottom strand is now predicted to code for p32, a precursor protein of 32.1 kDa (called p28K in Vrati *et al.,* 1996b) (Fig. 1A and Table 1). Inspection of the sequence showed several polyadenylation signals (AATAAA) 3′ to LH1, 2, and 3 (Table 1) that suggested possible transcript termination points. Our initial strategy therefore was to use the 5′ oligo AK1 (Table 1) together with a 3′ primer (AK7, 10, 11) from each of LH1–3 to ascertain whether cDNA products could be amplified by PCR.

LH1 transcripts

With primers AK1/AK7, two products of ~750 and 550 bp were amplified after 40 cycles of PCR from infected-cell RNA harvested at 24 h pi (Fig. 1B, lane 5). Sequencing revealed that they were derived from related, differentially spliced RNAs that were produced from a promoter (LHP1) upstream of base 427. The larger transcript, LHP1.1, contained sequences that were spliced out in LHP1.2 (Fig. 1A and Table 1). Both transcripts contained only one significant ORF that could code for the 9.7-kDa product. Termination of these RNAs probably occurred at possible polyadenylation sites at position 1390–1395 or 1467–1471 (Fig. 1A and Table 1) and generally did not extend into ORF LH2 because little or no such PCR product was detected using AK1/AK10 or AK1/AK11 primer combinations (data not shown) and because RACE–PCR with anchor/AK10 did not amplify a product greater than ~525 bp (Fig. 1C, lane 6) (that corresponded to LHP2.3, see below). However, AK1/AK10 or AK11 primers amplified a products of ~600 and 800 bp, respectively. A clone from the latter corresponded to LHP1.3 in which splicing occurred between position 815 and a new site at position 1464, i.e., downstream of the polyadenylation sites (Fig. 1A, Table 1).

To map the cap sites for LHP1.1 and LHP1.2, RACE–PCR was carried out using the 5′ anchor oligo together with 3′ AK7. Major and minor products of ~450 and ~600 bp, respectively, were obtained (Fig. 1C, lane 2) and the former was cloned. The product at ~170 bp was considered too short to be a significant transcript. Sequencing identified clones that carried the expected 5′ (anchor oligo) and 3′ terminal (AK7) sequences and confirmed splice junctions in LHP1.1 and LHP1.2. Unexpectedly, however, the 5′ ends of these clones mapped to positions 545 and 548 in the genome, i.e., substantially downstream of the proposed LHP1 promoter region, suggesting the presence of a second promoter, LHP2 (Fig. 1A and Table 1). It is possible that the ~600-bp product was derived from transcripts that originated at the LHP1 promoter.

Temporal analysis of LHP1 and LHP2. To confirm the presence of two promoters temporal transcription was investigated semiquantitatively with separate primer pairs in independent PCR reactions using the same cDNA prepared from polyadenylated mRNA harvested at 12, 24, and 48 h pi and the same number of PCR cycles. RNA synthesis from the LHP1 promoter was monitored by PCR using primers AK1/AK7, a combination that excluded transcripts from the LHP2 promoter because AK1 was upstream of the LHP2 start sites (Fig. 1A). Products of ~750 and 550 bp representing LHP1.1 and LHP1.2 transcripts were weakly detectable at 12 and 48 h pi in some experiments (data not shown) and readily detectable at 24 h pi (Fig. 1B, lane 5). Because transcripts derived from the LHP2 promoter shared sequences in common with transcripts derived from LHP1, LHP2 transcription could not be monitored exclusively by PCR. Transcription from both promoters was investigated using AK3/AK7 primers. An abundant product of ~380 nucleotides was amplified from infected-cell RNA (Fig. 1D, lanes 3, 5, and 7). The identity of this product was confirmed by digestion with *Hpa*II, which produced the expected bands of 200 and 157 bp (data not shown). That the PCR profile was temporally very different from the AK1/AK7 profile (Fig. 1B) confirmed the existence of two promoters in the region. The identification of a single PCR product (Fig. 1D, lane 3) also suggested that only the shorter LHP2.2 RNA was produced from LHP2 promoter. A similar temporal profile of LHP2 transcription was obtained with the AK3/AK10 primer pair (data not shown).

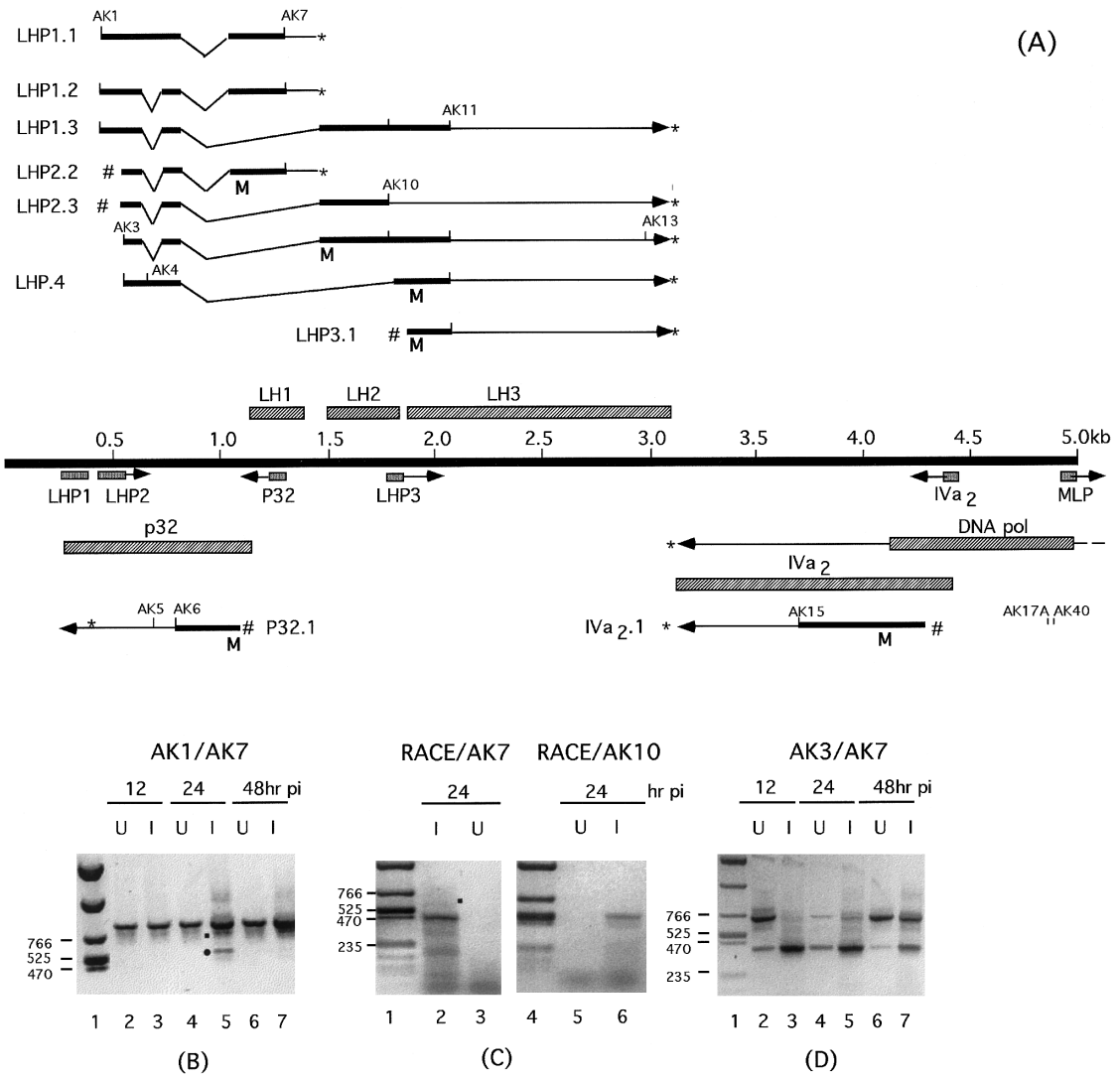

FIG. 1 Transcription map of the OAV lefthand end and IVa_2 regions. cDNA harvested from uninfected (U) and OAV-infected (I) CSL503 cells at 12, 24, and 48 h pi was reverse transcribed into cDNA using oligo$(dT)_{15}$ or random primers. Portions of selected RNAs were amplified by PCR or RACE–PCR using oligonucleotide primers (e.g., AK1, AK7 etc.) as described in the text. PCR products were cloned and sequenced to confirm their identity and to determine the location of splice junctions. In (A) the relevant portion of the genome is indicated. Predicted ORFs (striped boxes) are shown above and below the line for the left to right and right to left orientations, respectively. Thick lines represent sequenced regions and thin lines indicate the likely extent of the transcript terminating with a polyadenylation signal (*). (#) Signifies that transcription start sites were mapped by RACE–PCR. (M) Indicates the first available Met in the ORF. Transcripts were named according to the promoter from which they were derived and numbered separately if their splicing pattern differed. The approximate location of promoters LHP1, 2, 3 etc. is indicated by shaded boxes. Arrows indicate the direction of transcription. The lower panels show PCR analyses that were carried out using the primer pairs indicated. Temporal transcription from the LHP1 and LHP2 promoters was analysed in B and D. Products that were cloned and sequenced in B and C are marked with a dot. (<) Signifies a possible LHP1 RACE–PCR product. (^) Marks the ~380-bp AK3/AK7 product. Inverted images of ethidum bromide-stained agarose gels are shown. The sizes of marker fragments are indicated.

LH2 and LH3 transcripts

To identify the 5′ ends of transcripts corresponding to LH2 (14.7 kDa) and LH3 (42.8 kDa), RACE–PCR was performed using the 5′ anchor oligo together with AK10 or AK11. With anchor/AK10 a product of ~525 bp was obtained (Fig. 1C, lane 6). With anchor/AK11, a product of ~750 bp as well as a shorter product of ~230 bp was obtained (data not shown). Cloning and sequencing identified the LH2.3 transcript that contained a new splice junction at position 1464 and shared upstream splice junctions with LHP11 and LHP12 RNAs (Fig. 1A and Table 1). The 5′ ends of five LH2 clones mapped to positions 540 (three clones), 552, and 567, indicating that they were also derived from promoter LHP2. Splice junctions in the LHP2.3 family were confirmed by analysis of AK3/AK10 PCR clones. The 5′ ends of the 230-bp (AK11) clones that contained the appropriate anchor oligo mapped to positions 1869, 1874, and 1891, indicating the presence of a third promoter, LHP3, that produced LHP3.1 transcripts (Fig. 1A and Table 1). These RNAs

TABLE 1

Summary of Lefthand End and IVa_2 Transcripts

ORF	RNA	Start sites	5′ Exon	Exon	Exon	Coding region	Poly(A) signal	Amplified by
LH1 9.7 kDa	LHP1.1 (1)[a]	ND	<427-815[b]	1061>	—	1143–1391	1390–1395 or 1467–1471	AK1/AK7[c] 427+/1318−
LH1	LHP1.2 (1)	ND	<427-634	744-815	1061>[b]	As above	As above	AK1/AK7
LH2 14.7 kDa	LHP1.3 (1)	ND	<427-634	744-815	1464>	1489–1872	3054–3059	AK1/AK11
LH1	LHP2.2 (2)	545, 548	545-634	744-815	1061>	1143–1391	1390–1395 or 1467–1471	RACE/AK7
LH2 14.7 kDa (E1B 19 kDa)[e]	LHP2.3 (8)	538, 552, (3)[d] (1) 575 (1)	538-634	744-815	1464>	1489–1872	3054–3059	RACE/AK10 1829+ and AK3/AK11 585+/2056−
	LHP.4 (1)	ND	<585-815	1828>		1915–3063	3054–3059	AK3/AK11
LH3 42.8 kDa (E1B 55 kDa)[e]	LHP3.1 (3)	1869, 1874, 1891	—	—	—	1915–3063	3054–3059	RACE/AK11
p32 32.1 kDa	P32.1 (2)	1144, 1161	—	—	—	238–1098 (− strand)	377–382?	AK5/RACE 723+
IVa_2 37.5 kDa	IVa_2.1 (2)	4268, 4273	—	—	—	3071–4055 (− strand)	3071–3076	AK15/RACE 3729+

[a] Number of clones analysed to determine splice junctions.
[b] (<) and (>) signifies that exon begins before or continues beyond the indicated nucleotide, respectively.
[c] The 5′ nucleotide and (+) or (−) sense of the primer is shown at first mention.
[d] Number of clones analysed to determine start sites.
[e] Shows homology with the Ad5 protein (Vrati *et al.,* 1996b).

were unspliced and probably terminated following the polyadenylation site at the end of LH3.

Because a possible polyadenylation site was present in LH2 at position 1860–1865, PCR was carried out with AK3/AK11 primers to ascertain whether LHP2.3 transcripts terminated at this point or continued into LH3 (Fig. 1A). A product of ~700 bp was obtained (data not shown) and sequenced, showing that it was a spliced transcript and confirming that LHP2.3 transcripts were extended (Fig. 1A). Similarly, primers AK3/AK13 produced a PCR product of ~1.6 kb (data not shown), indicating that LHP2.3 transcripts also terminated at the common polyadenylation site at 3054–3059, as previously mapped (Vrati *et al.,* 1996a). Using AK3/AK11 primers a single PCR clone (LHP.4; Fig. 1A) was also obtained in which the intron between bases 635 and 743 was retained, while a longer intron between nucleotides 815 and 1828 was spliced out. In an attempt to specifically amplify this transcript from infected-cell RNA, primer AK4, from within the retained intron, was used together with AK11 (Fig. 1A). However, this combination did not amplify the expected product of ~400 bp but did produce a product of ~1.4 kb, expected from residual genomic DNA in the 48-h sample (data not shown).

p32k transcripts

The OAV genome also codes for a proteolytically processed structural protein on the bottom strand at the LHE (Vrati *et al.,* 1996b). The revised OAV sequence (GenBank Accession No. U40839) and peptide data (Vrati *et al.,* 1996b) predicts that this is a 30.9-kDa protein derived from a 32.1-kDa precursor (Table 1). To locate the P32 promoter, RACE–PCR was performed using 5′ anchor/AK5 primers (Fig. 1A). A virus-specific product of ~340 bp was obtained by nested PCR with AK6. Two clones of appropriate size that carried the 5′ anchor oligo were identified by sequencing. Neither clone contained an intron and the 5′ ends of these P32.1 transcripts mapped to positions 1144 and 1161 on the negative strand (Fig. 1A and Table 1), suggesting that the P32 promoter was nearby. No p32k-like clones were obtained when AK3/AK7 primers were used for PCR, suggesting that the P32.1 transcript initiated to the left of the AK7 sequence. A product of 733 bp should have resulted from amplification of unspliced p32k transcripts or residual genomic DNA. Some material of this size was produced (Fig. 1D), but it was not virus specific. Two putative p32 clones of this size did not contain OAV sequences. The first initiation codon in ORF p32 occurred 12 amino acids upstream of the previously identified protease cleavage site (Vrati *et al.,* 1996b).

IVa2 transcripts

The location of the IVa_2 promoter was also investigated. Because the IVa_2 promoter lies adjacent to the MLP in Ad2 (Natarajan *et al.,* 1985) it was thought that a similar situation might exist in OAV. Primers were designed from an analysis of conserved sequences and possible initiation codons in the vicinity of the MLP. PCR performed with primers AK15/AK17A (bases 4744–4768)

and AK15/AK40 (bases 4793–4819) (Fig. 1A) amplified a virus-specific product of ~400 bp that was cloned and sequenced. Two clones that were derived from spliced RNAs were identified (data not shown) but they did not represent IVa_2 transcripts. They had a common splice site at base 3791, well within the IVa_2 ORF, different rightward splice junctions at positions 4488 and 4515, and a reading frame that was not continuous with the predicted IVa_2 ORF. In fact the clones identified frame-shifted transcripts that potentially would have coded for DNA pol with C-terminal truncations. RACE–PCR was then carried out using the AK15/5′ anchor primer combination and the same anchor-modified RNA preparation that was used to map the E4 and DBP transcription start sites (see below). The largest product obtained was ~360 bp in size. Two clones were obtained in which the 5′ ends mapped to positions 4268 and 4273 (Table 1). Use of the first available initiation codon 74 residues downstream would produce a protein of 37.5 kDa.

E2 genes and transcripts

OAV E2 genes coding for DNA binding protein (DBP), terminal protein (TP), and DNA polymerase (DNA pol) were clearly identified by their homology with other Ad genes (Vrati *et al.,* 1996a, b). For Ad2, expression of the E2 proteins is driven by early promoters E2A and E2B. These are located between the 33K and pVIII genes and within the 100K hexon assembly gene, respectively (Horwitz, 1990). Comparison of OAV and Ad5 nucleotide sequences suggested that OAV promoters might also be located in these positions.

DBP transcripts

To test this, 5′ primers AK31 and AK32 were used in conjunction with DBP primer AK28. The AK28/AK31 combination failed to amplify a product, but AK28/AK32 produced an abundant virus-specific band of ~850 bp from RNA harvested at 12, 24, or 48 h pi (see Fig. 5 of Khatri *et al.,* 1997). Direct sequencing of the product identified transcript DBP.1 and its exons. The data were confirmed by RACE–PCR and sequencing again using 5′ anchor/AK28. The 5′ end of a clone identified by this means mapped to position 21,380, eight bases upstream of AK32. Thus, DBP.1 has a 5′ exon and a second exon of 98 nucleotides upstream of the ORF (Fig. 2A, and Table 2) and the first Met was located within the DBP ORF (Fig. 2A and Table 2), predicting a protein of 43.5 kDa. DBP.1 transcripts probably terminate near the first available polyadenylation signal beyond the ORF at position 17895–17900.

Terminal protein (pTP) transcripts

pTP transcripts were identified by PCR with primers AK21a/AK32 (Table 2 and Fig. 2A). Virus-specific products of ~750 and ~250 bp were obtained but their ratios varied with the time pi. The larger product decreased in amount from 12 to 48 h pi, while the smaller one increased (Fig. 2B, lanes 3, 5, and 7). The larger product yielded clones of two types, TP.1 and TP.2, that differed in the number of exons they contained (Fig. 2A and Table 2). The 5′ exons were common, but the third exon in TP.1 was absent from TP.2. TP.1 exon 4 and the splice junction with the major ORF were again shared with TP.2. The initiation codon for pTP is not within the major ORF. Splicing out the last intron before the main ORF linked the amino acid sequence MTARQ, in the penultimate exon, in-frame to IHQWQQ at residue 65 of the major ORF (Fig. 2A).

Clone TP.3, derived from the short PCR product produced at late times pi (Fig. 2B, lanes 5 and 7), had a longer 5′ exon (Table 2). Splicing apparently occurred between the sequence TGGATATTTTT at position 21,263 on the minus strand to the same sequence (underlined) at position 8654 in the TP ORF, creating an mRNA with the sequence TGGATAGAATAT. The first available Met in the new transcript occurred a further 10 residues downstream of the junction and formed part of a potential protease cleavage site, i.e., MQGF'G. TP transcripts probably terminate near the polyadenylation signal at the end of the ORF at position 7344–7349 (Fig. 2A).

DNA pol transcripts

DNA pol transcripts were identified by PCR with primers AK19/AK32 (Fig. 2A). A virus-specific product of ~800 bp was amplified at 12 h pi. At 24 and 48 h many other virus-specific products were produced (data not shown), some of which appeared to be amplified from late RNAs. These obscured the DNA pol-specific product. The 800-bp product corresponding to DNA pol.1 contained three exons, all of which were common to TP.2 transcripts (Fig. 2A and Table 2). The initiating Met was within the major ORF. DNA pol transcripts probably terminate at the first available polyadenylation signal at the end of the ORF for IVa_2 at position 3071–3076.

E4 transcripts

The revised genome sequence (GenBank Accession No. U40838) now shows that the putative E4 region comprises three major ORFs, potentially coding for proteins of 17.1 kDa (E41), 25.6 kDa (E42), and 30.85 kDa (E43) (Fig. 3 and Table 3). Because the location of the promoter for this region was unknown, primers AK34, AK35, and AK36 from each of the ORFs (Fig. 3A and Table 3) were used in RACE–PCR analyses to map the 5′ ends of E4 transcripts. The largest virus-specific product that was detected with RACE/AK34 was ~300 bp in size (Fig. 3B, lane 3). This yielded two types of clones with the correct terminal anchor/primer sequences (P1E43.1 and P2E43.1). The clones shared downstream exons and splice junctions but differed in their 5′ exons with tran-

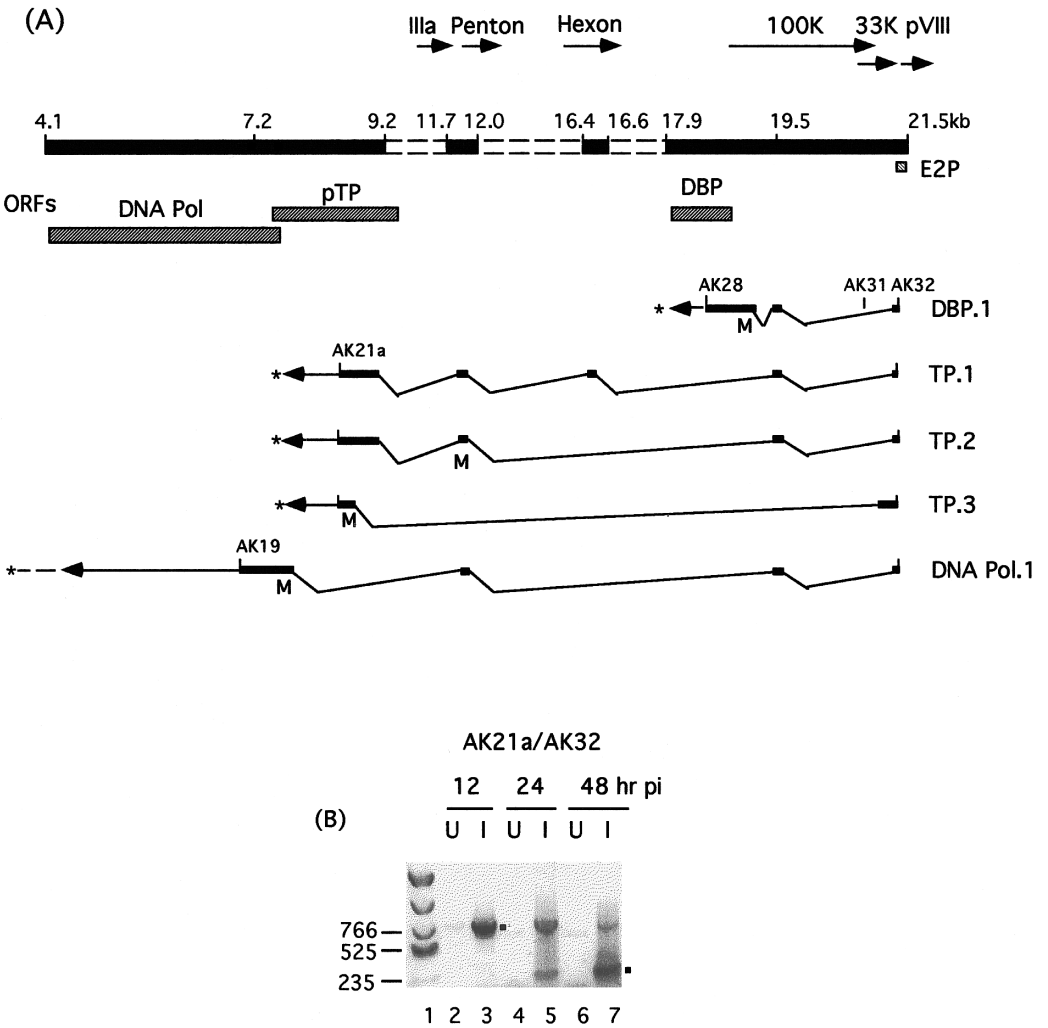

FIG. 2. Transcription map of the OAV E2 region. Details of nomenclature are as for Fig. 1 The E2 promoter and transcripts for DBP, pTP, and DNA pol are indicated. The probable initiation codon for pTP lies outside the main ORF. (B) Temporal PCR analysis of pTP transcripts using primers AK21a/AK32. The ~760- and <235-bp fragments yielded clones corresponding to transcripts TP1/2 and TP.3, respectively.

scription start sites being located at positions 26,030 and 26,251 (Fig. 3A and Table 3). This suggested the presence of two separate promoters, P2E4 and P1E4, respectively, within the ~1-kb A/T-rich region of the genome that was previously of unknown function (Vrati *et al.,* 1996b). RACE–PCR amplification with primers AK36 and AK35 produced virus-specific products of ~420 and ~520 bp, respectively (Fig. 3B, lanes 6 and 8). However, only clones with 5′ ends that mapped between positions 25,985 and 26,039, i.e., to the P2E4 promoter, were obtained, e.g., P2E411 and P2E421 (Fig. 3A and Table 3).

Temporal analysis of E4 promoters

To confirm the existence of two promoters, new primers from within the 5′ exons were made (AK37 and AK39; Table 3 and Fig. 3A) and used for PCR in conjunction with the downstream primers AK33, AK35, and AK36, using the same cDNA preparations and the same number of PCR cycles. Products were amplified from P2E4 RNAs harvested at 12, 24, or 48 h pi with AK33/AK37 (Fig. 3C, lanes 3 and 5 and data not shown), AK35/AK37, and AK36/AK37 (data not shown). In general, these clones had the same exons and splice junctions as the clones above isolated by RACE–PCR, although some variations were noticed. For example, in P2E412 the intron between exons 3 and 4 was retained (Fig. 3A and Table 3). However, no detectable virus-specific products were amplified when the P1E4 primer AK39 was used with E41 and E42 primers AK35 and AK36 (data not shown), suggesting that E41 and E42 RNAs were not expressed from the P1E4 promoter and that the P1E4 promoter may only function at early times to generate E43 transcripts. In contrast, PCR with primers AK33/AK39 generated products of ~500 and ~700 bp from RNA harvested at 12 h pi, but not at later times (Fig. 3D, lanes 3 and 5). (The product of <235 bp (lane 3) was considered too short to

TABLE 2

Summary of E2 Transcripts

ORF	RNA	Poly(A) signal	Coding region	Exon	Exon	Exon	Exon	5′ Exon	Amplified by
DBP 43.5 kDa	DBP.1[a]	17895–17900	17964–19113	<19142	—	—	19460–19558	21318->21372 21380[c]	AK28/AK32[b] 18301+/21372− AK/28/RACE
pTP 69.8 kDa	TP.1 (3)[d] (early) TP.2 (1)	7344–7349	7348–9120	<9120 <9120	11881–11989 11881–11989	16448–16510	19460–19558 19460–19558	21318->21372 21318->21372	AK21a/AK32 8615+
TP 49.6 kDa	TP.3 (4) (late)		7348–8619	<8647	—	—	—	21258->21372	
DNA pol 125.34 kDa	DNA pol.1 (1)	3071–3076	4151–7365	<7574	11881–11989	—	19460–19558	21318->21372	AK19/AK32 6958+

[a] PCR product initially sequenced directly, i.e., not cloned.
[b] See Table 1, note C.
[c] DBP.1 start site determined for one clone.
[d] Clones sequenced to determine splice junctions.

be significant given the location of the AK33 primer at the 3′ end of ORF E43.) A fragment of ~2000 bp, probably amplified from residual viral DNA, was observed at 48 h pi (Fig. 3D, lane 7). The 5′ exons of P1E43.2 and P1E43.3 were differentially spliced, although they shared a downstream exon and splice junctions (Fig. 3A and Table 3). The temporal comparison of P1E4 and P2E4 promoter function using AK33/AK37 and AK33/AK39 primer pairs showed clearly that P1E4 functioned only at early times while P2E4 also functioned at later times (Figs. 3C and 3D).

Early transcripts from the MLP

During Ad2 infection, the major late promoter is active at early times but the splicing pattern of early L1 transcripts differs from that of late transcripts which have the characteristic tripartite leader sequence (TLS) (reviewed in (Horwitz, 1990). To ascertain whether a similar change in splicing pattern occurred during OAV infection, PCR was performed using a primer from within TLS exon 1 (AK18) and AK23, located in the OAV 38.2-kDa ORF (Fig. 4 and Table 4). PCR products amplified from RNA harvested at 12, 24, and 48 h pi differed greatly. A minor product of ~1400 bp was identified at 12 and 24 h pi, while a major product of ~500 bp was present at 24 h and to a lesser extent at 48 h pi (Fig. 4B, lanes 3, 5, and 7). An additional minor product of ~800 bp was also detected at 24 h pi. Clones of the ~1400- and ~500-bp fragments were obtained and sequenced. Most of the 500-bp clones were derived from MLP L12 and contained the previously described OAV TLS (Vrati *et al.,* 1996a), but a variant (MLP L13) that lacked exon 2 was also identified (Fig. 4, Table 4). The 1400-bp fragment was derived from MLP L11 transcripts that lacked an intron following what is exon 3 in L12. The initiation codon of the 38.2-kDa ORF was unchanged.

DISCUSSION

OAV is distinguishable from the Mast- and Aviadenoviruses by several criteria. Phylogenetic analysis of the protease and hexon genes groups OAV with BAV-7 and EDS viruses (Harrach *et al.,* 1997; Hess *et al.,* 1997). In addition, the distinctive genome arrangement and high A/T content of OAV, the presence of at least one unique viral capsid protein and the absence of others, and the characteristics of its nonstructural proteins (Vrati *et al.,* 1995; Vrati *et al.,* 1996b; Xu *et al.,* 1997) separate OAV from other genuses of the Adenoviridae. This was recently confirmed by analysis of the related EDS genome (Hess *et al.,* 1997). Therefore, the location of promoters controlling transcription units in OAV could not be readily deduced by comparisons with Mastadenoviruses, the only viruses for which data were available (reviewed in (Horwitz, 1990). Some data describing late transcripts and early RNAs from the RHE of OAV were obtained

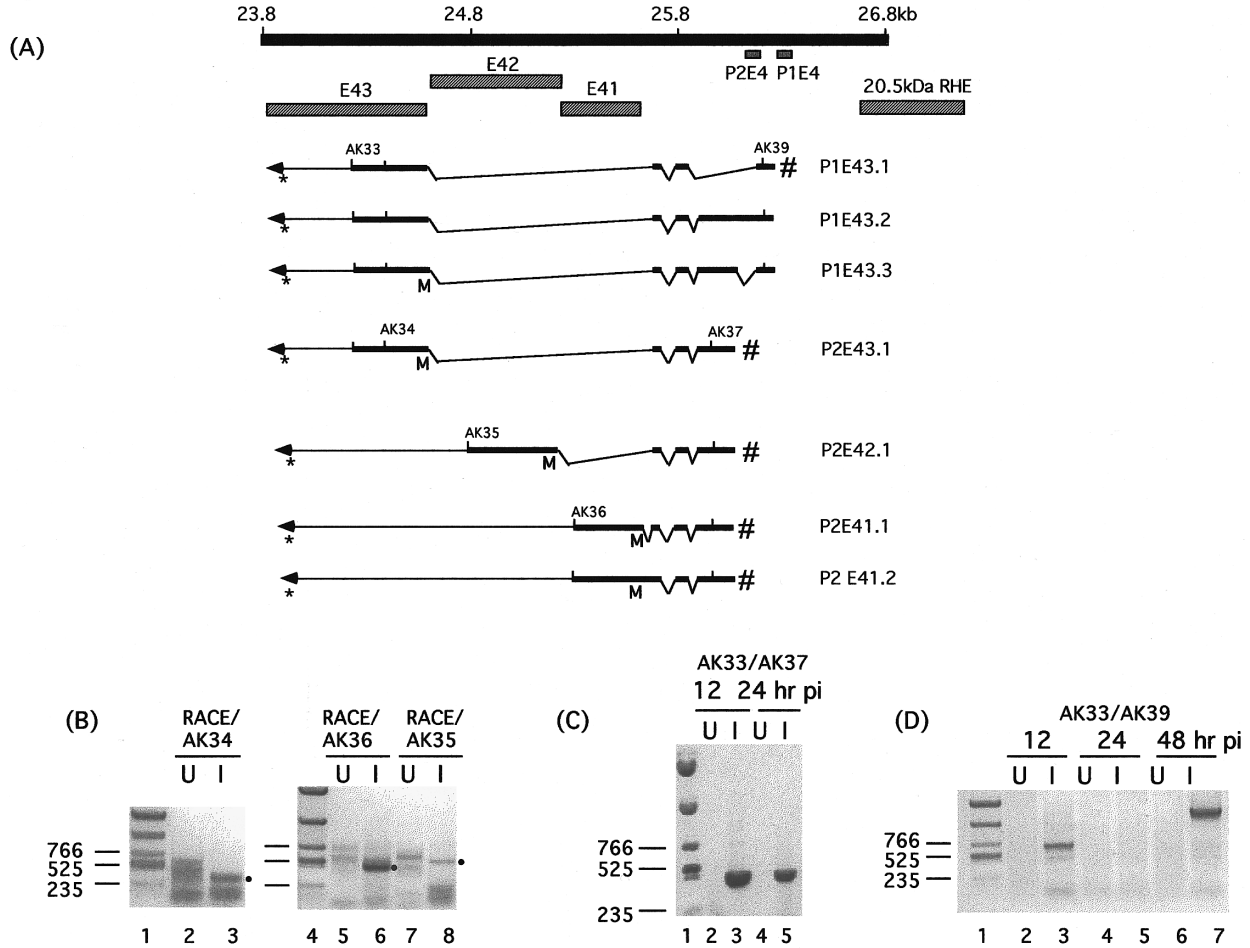

FIG. 3. Transcription map of the OAV E4 region. The nomenclature is described in Fig. 1. Promoters P1E4 and P2E4 and the transcripts derived from them are indicated. RACE–PCR (B) and temporal analysis of selected P2E4 (C) and P1E4 (D) transcripts were carried out using the primer pairs indicated.

previously (Vrati *et al.,* 1995, 1996a; Xu *et al.,* 1997). To gain an overview of transcription in OAV, a PCR-based strategy was used to identify early transcripts and their splice junctions as well as the 5′ ends of RNAs. This achieved a better definition of probable reading frames in the genome and defined the likely location of viral promoters.

Several points were considered in relation to the PCR strategy. Firstly, the 5′ ends of transcripts were mapped by RLM–RACE and RT–PCR in an attempt to identify the locations of promoters in the genome. This procedure was selected to minimise the detection of truncated RNAs. Only capped transcripts that retained a 5′ phosphate after alkaline phosphatase/pyrophosphatase treatment were capable of accepting the anchor oligonucleotide during ligation (Maruyama and Sugano, 1994). Only in the event that the 5′ end of an RNA was unavailable for ligation would a species not be represented in the population. As a further safeguard against truncation, only full-length cDNA molecules in which the anchor sequence was copied by reverse transcriptase could then be amplified by PCR using the anchor oligo in combination with a downstream primer. The discrete nature of RACE–PCR products that we obtained coupled with RT–PCR analyses of GAPDH in the RNA samples (see Materials and Methods) argues that the transcripts analysed in these studies were intact. Secondly, to overcome the imperfect specificity of primer pairs in some PCR reactions, apparent virus-specific products were cloned and positively identified by nucleotide sequencing. Comparisons against the known genomic sequence of OAV (Vrati *et al.,* 1995, 1996a, b) identified splice junctions within transcripts. PCR profiles were used only to demonstrate temporal differences in promoter function. Thirdly, PCR amplification may not have resulted in the quantitative detection of all potential products. Single primer pairs were used in all PCR reactions to amplify families of related transcripts. However, the data may not reflect the relative abundance of transcripts from a given promoter, nor were all differentially spliced transcripts necessarily identified. Lastly, the location of transcription start sites identified for LHP2 and E4P2 transcripts were

TABLE 3

Summary of E4 Region Transcripts

ORF	Promoter transcript	Coding region	Exon	Exon	Exon	Exon	5′ Exon	Amplified by
E43 30.85 kDa (E4/34K protein)[a]	P1E43.1	23851–24564	<24593	25731–25752	25838–25878		26213-26251[b] (1)[c]	AK34-RACE[d] 24419+ AK33/AK39 24220+/26237−
	P1E43.2		<24593	25731–25752	25838–25878		25965->26237 (1)	AK33/AK39
	P1E43.3		<24593	25731–25752	25838–25878	25965–26043	26213->26237 (1)	AK33/AK39
	P2E43.1		<24593	25731–25752	25838–25878		25965-26030 (1)	AK34-RACE
E42 25.6 kDa (E4/34K protein)[a]	P2E42.1	24561–25233	<25238	25731–25752	25838–25878		25965-25933, 26039 (4)	AK35-RACE 24826+ AK35/AK37 25989-
E41 17.1 kDa	P2E41.1	25223–25651	<25651	25731–25752	25838–25878		25965-25985 (2)	AK36-RACE 25325+ AK36/AK37
	P2E41.2			<25651–25752	25838–25878		25965-25993 (5)	AK36-RACE AK36/AK37

[a] Shows limited homology with Ad2 ORF6 34-kDa protein (Vrati *et al.,* 1996b).
[b] Underlined numbers signify one or more cap sites determined.
[c] Clones analysed to determine splice junctions.
[d] See Table 1, note *C*.

more widespread than observed for TATA-element containing Ad2 promoters in an earlier study (Baker and Ziff, 1981). However, with improved cDNA cloning procedures similar to those used in our work a wider distribution of cap sites has been found for a large variety of transcripts, particularly when polypyrimidine tracts occurred just upstream (Kato *et al.,* 1994). Polypyrimidine tracts that are located 5′ proximal the cap sites for LHP2 and E4P2 promoters may influence the initiation of transcription.

Transcription units in the genome

There appear to be seven transcription units in the OAV genome controlled by 10 promoters. For the P32, IVa_2, E2, and RHE units individual promoters appear to control expression. The surprising observation was that for the LHE three promoters appear to control the transcription of three ORFs. The LHP1 promoter appears to be genuine, rather than an infrequent, alternative producer of related transcripts (Osborne *et al.,* 1982) because it functioned in a temporally distinct manner. Similarly, dual E4 promoters apparently exist, as defined by the identification of RACE–PCR transcripts with common downstream sequences and distinct 5′ exons. These promoters also appeared to operate at different times during infection in permissive CSL503 cells as identified by their temporally distinct PCR profiles (Table 5). Although the PCR analyses were not quantitative, at a gross level, based on amplification of the same cDNA preparation for the same number of cycles, the LHP2 and P2E4 promoters were more active than the LHP1 and E4P1 promoters, respectively.

LHE transcription

The LHE of the genome potentially codes for three proteins on the top strand and one structural protein on the bottom strand (Fig. 1A). The OAV LH2 and LH3 ORFs show limited homology with the E1B proteins of Ad2 (19 and 20% identity for LH2 and LH3, respectively, when aligned by standard GCG parameters described in Vrati *et al.,* (1996b) but are closer to their EDS homologues (29 and 35% identity, respectively). Similarly, the p32 polypeptide has an EDS counterpart (Vrati *et al.,* 1996b; Hess *et al.,* 1997). LH1 shows no detectable homology with any other Ad. RNAs coding for the LH1 product were transcribed from either the LHP1 or LHP2 promoters but appeared not to extend beyond the polyadenylation signals at the end of this ORF as judged by the RACE–PCR products (Fig. 1C) and because AK1/AK10 or AK11 primer combinations failed to produce a PCR product of a size consistent with a continuous transcript. Similarly, transcripts for LH2 were detected from both promoters, and these probably continued through ORF LH3. LHP3.1 RNAs apparently originated from the LHP3 promoter located between the LH2 and LH3 ORFs, because cap sites that mapped between positions 1870 and 1892 were identified by RACE–PCR. There is no strong candidate TATA box upstream of these start sites.

A clone that represented transcript LHP.4 (Fig. 1A), derived from the LHP1 or LHP2 promoters, was amplified with primers AK3 and AK11. Potentially, this transcript could have represented the major transcript for LH3, spliced so that the first Met in the LH3 was the initiation codon. This would have removed the need for an LHP3 promoter. However, LHP.4 contained a splice junction at position 1828, a site that was not predicted from the OAV

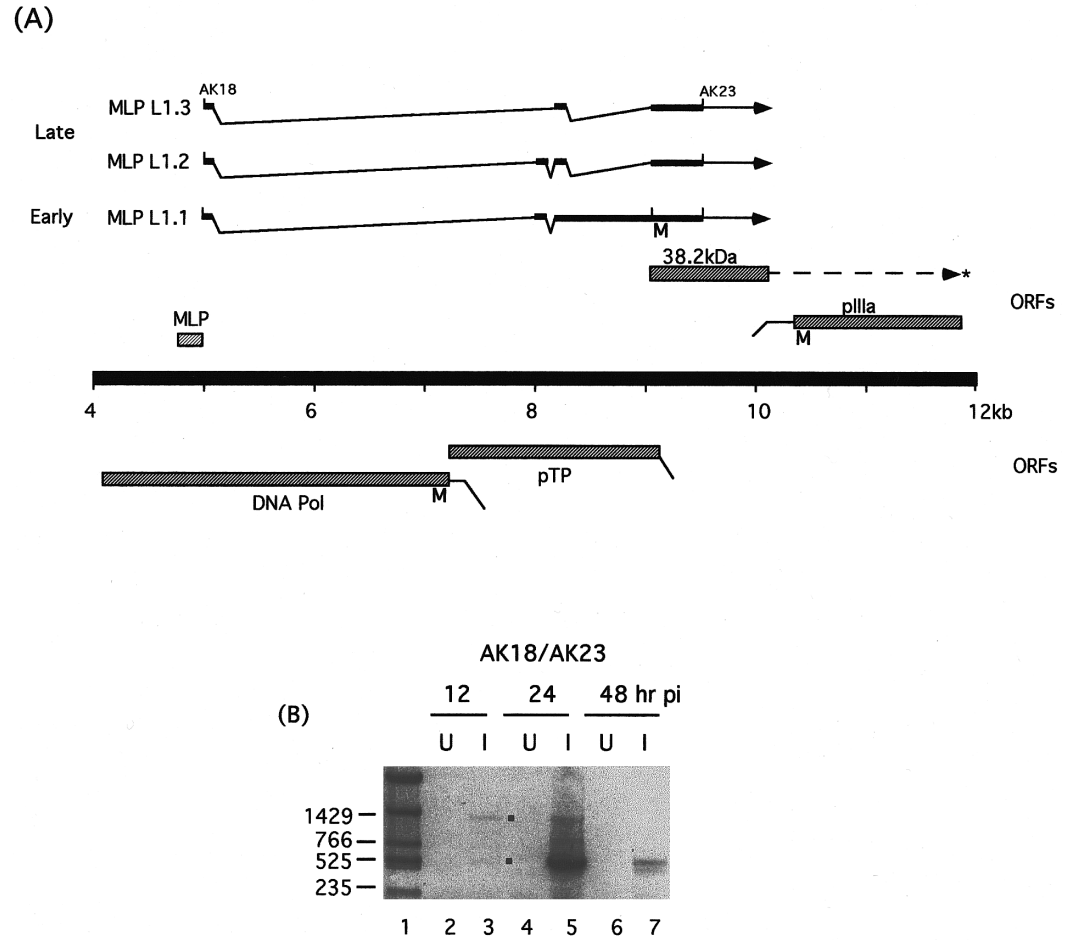

FIG. 4. Transcripts and leader sequences from the major late promoter. In A the differentially spliced transcripts coding for the 38.2-kDa protein are shown. In (B) a PCR analysis using primers AK18/AK23 produced virus-specific products of ~1400 bp (lane 3) and ~500 bp (lanes 3 and 5). These were cloned and sequenced to identify the early and late leader sequences shown in (A).

genome sequence because an AC′GT, rather than an AG′GT sequence was present. LHP.4 also retained an intron between bases 635 and 743 (Fig. 1A). Using sequences in this intron, PCR amplification was performed to selectively amplify the transcript. However, no product of the expected size was obtained. Therefore it is concluded that the LHP.4 clone represented an atypical transcript and that LH3 RNAs are derived from LHP3, rather than an upstream promoter.

The transcription map (Fig. 1A) derived by the PCR strategy agreed with an analysis of LHE RNAs by Northern hybridisation. Using a probe that spanned the first

TABLE 4

Summary of MLP Transcripts

ORF	RNA	5′ Exon	Exon	Exon	Exon	Coding region	Poly(A) signal	Amplified by
38.2 kDa (52/55K)[a]	L1.1 (2)[b]	5013–5133	8082–8144	8349–9158>		9158–10165	11862–11867	AK18/AK23[c] 5013+/9522−
	L1.2 (4)	5013–5133	8082–8144	8349–8411	9158>			AK18/AK23
	L1.3 (1)	5013–5133	—	8349–8411	9158>			AK18/AK23
pIIIa 58.4 kDa					10099>	10316–11863	11862–11867	Vrati *et al.* (1996a)

[a] Shows homology with the Ad5 protein (Vrati *et al.,* 1996a,b).
[b] Number of clones sequenced to determine splice junctions (see also Vrati *et al.,* 1996a).
[c] See Table 1.

TABLE 5

Summary of OAV Promoter Activity

Promoter	Time pi (h)		
	12	24	48
LHP1	−	+	+
LHP2	+	+	+
LHP3[a]	−	+	nd
E2[b]	+	+	+
P1E4	+	−	−
P2E4	+	+	+
RHE[b]	+	+	+

[a] Based on Northern analysis (Z. Xu, unpublished results).
[b] Khatri *et al.* (1997).

2.5 kb of the genome, RNAs of ~0.6 kb were detectable as early as 8 h pi and at 24 h pi additional species of ~1.4 and ~2 kb were present (Z. Z. Xu, unpublished results). These sizes agree well with those predicted from Fig. 1A. In particular, because the LHP3.1 transcript was expected to be ~1.4 kb in size (Fig. 1A), its appearance at a time distinct from the smaller LHP2 products confirmed the existence of the LHP3 promoter.

It is also of interest that LHP2.3 RNAs continued into the LH3 ORF, probably terminating at a common point downstream (Fig. 1A). Such an arrangement is consistent with transcription of the Ad5 genome where readthrough transcription from the upstream E1A region provided a *cis*-acting effect that relieved repression of the E1B promoter prior to DNA replication (Maxfield and Spector, 1997). The structural protein p32K also has its own promoter and produces an unspliced transcript. In both senses it is similar to protein IX production by Mastadenoviruses.

IVa$_2$ promoter

Because the IVa$_2$ promoter is adjacent to the MLP in Ad2 (Natarajan *et al.,* 1984), attempts were also made to identify upstream transcription start sites by comparing OAV MLP sequences with Ad2. The region also codes for the highly conserved DNA pol, which made it possible to deduce the theoretical position of the IVa$_2$ transcription start site in OAV and identify a nearby AT-rich motif TATAAAAA (Chen *et al.,* 1994). However, PCR products amplified with AK15 and two primers from this region did not identify spliced RNAs that were linked to the IVa$_2$ reading frame, although clones that potentially modified the C-terminus of DNA pol were identified. The latter most probably represented unproductive splicing events similar to others identified for pVIII RNAs in previous work (Vrati *et al.,* 1995), although analysis of DNA pol itself is required to determine whether the enzyme is modified during infection.

The IVa$_2$ promoter region was subsequently identified by RACE–PCR. The transcription start sites for the largest product mapped ~130 bases into the ORF. The results are supported by the fact that the start sites were identified using the same anchor-modified RNA preparation that was used to map the LHP2 and E4 start sites and promoters. If the location of the IVa$_2$ promoter is confirmed by other means, its location would suggest that the IVa$_2$ transcripts are unspliced. The initiation codon located ~74 residues downstream of the start sites was the first available in the ORF. Initiation at this point would produce a protein of 37.5 kDa. This is shorter than the theoretical IVa$_2$ protein of EDS which has a Met codon that is absent from OAV (Vrati *et al.,* 1996a; Hess *et al.,* 1997). The next available Met in EDS occurs past the end of the overlapping DNA pol ORF near the proposed initiation site for OAV and would produce a protein of 36.1 kDa. The N-terminal portions of IVa$_2$ from CELO, EDS/OAV, and Ad2/Ad40, representing the three genera, differ in length and show little homology compared with downstream sequences from residue ~150 of Ad2 to the C-terminus, which are substantially conserved. However, the shorter EDS and IVa$_2$ proteins contain all the conserved sequences (Fig. 5).

Control of E2 transcripts

For the E2 transcription unit there was a distinct difference in control between the Mastadenoviruses and OAV. During Ad2 infection DBP expression at early times is controlled by the E2A promoter, located between the 33K and pVIII ORFs, while at later times DBP, pTP, and DNA pol are expressed from the E2B promoter in the 100K ORF. Thus, TP and DNA pol are not expressed from E2A (reviewed in (Horwitz, 1990; Shenk, 1996). For OAV, a single promoter located between the 33K and pVIII ORFs appeared to control expression of the whole region because we identified transcripts for pTP and DNA pol that were derived from the same promoter as DBP. The initiating methionine for DBP and DNA pol apparently resides within the respective major ORF, but for pTP it was located in the penultimate exon (Fig. 2A). Thus, the TP coding region (69.8 kDa) is shorter than was previously thought (Vrati *et al.,* 1996a, b).

A distinct change in the processing pattern for pTP transcripts was also observed during infection, suggesting that alternative RNA splicing and processing may be used to produce a shorter form of pTP later in the infectious cycle. A splicing event that linked an extended 5′ exon directly to residue 220 in the major ORF eliminated upstream methionines and positioned Met$_{231}$ as the first initiation codon, suggesting the production of a truncated form of pTP with the N-terminal sequence MQGF′G. This sequence is also a site for processing of pTP to iTP by the viral protease (Webster *et al.,* 1994). However, the sequence across the splice junction of these clones (TGGATAGAAT; 5′ exon sequence underlined) did not

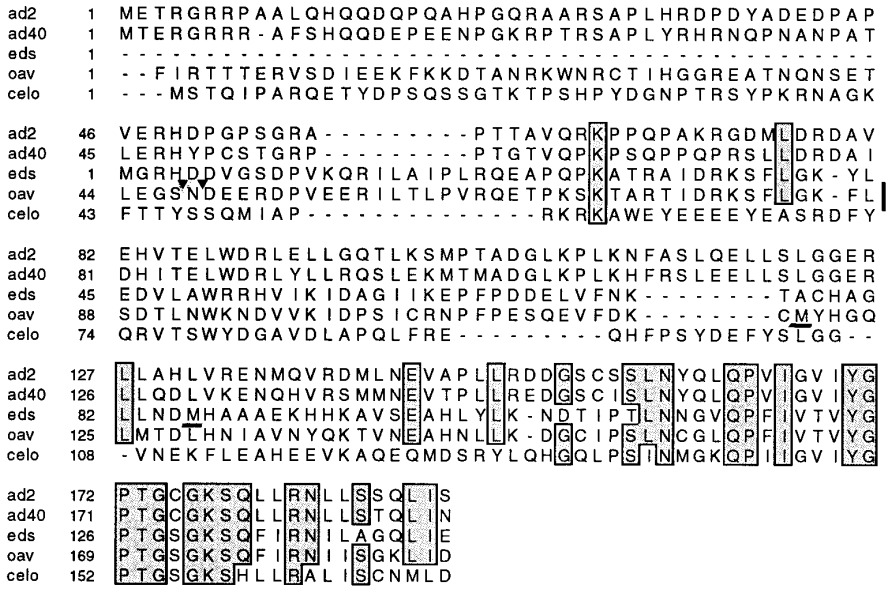

FIG. 5. Comparison of the amino-terminal regions of IVa_2 proteins. Sequences for Ad2, Ad40, CELO, EDS, and OAV are compared. Sequence identity between the genera is indicated by the boxed regions. A similar degree of identity extends to the C-terminus of the proteins. In the OAV sequence, triangles indicate the location of cap sites, the vertical bar shows the end of DNA pol in the overlapping ORF. Potential initiating methionine residues downstream of the DNA pol ORF are underlined.

conform to major (GT-AG) or minor (AT-AC) splice sequences (Kreivi and Lamond, 1996). Thus, the derivation of this transcript is unclear.

E4 region transcription

E4 transcripts with distinct 5′ exons identified two promoters that lie within ~1 kb of A/T-rich sequences whose function was previously unknown (Vrati *et al.,* 1996b). The promoters may have different levels of activity, with P2E4 > P1E4. The P2E4 promoter was active at 12 h and throughout infection (Table 5) and produced transcripts for all three E4 ORFs. In contrast, P1E4 activity was detectable at 12 h but not at 24 or 48 h pi and only transcripts that coded for ORF E43 appeared to be produced. This ORF potentially codes for a protein that has some homology to the *Mastadenovirus* E4 34-kDa protein (Vrati *et al.,* 1996b). Further work is required to ascertain whether the P1E4 promoter is one of the earliest active promoters during infection.

L1 transcripts from the MLP

Products transcribed from the OAV MLP at early times were clearly distinct from those produced late in infection. At 12 h a much longer transcript was produced in which the intron separating TLS exon 3 from the 38.2-kDa ORF was retained. By 24 h pi, when late transcription was in progress, the dominant product was the shorter RNA that carried the TLS (Vrati *et al.,* 1996a) or a variant of it (Fig. 4). The longer transcript contains a leader that may be the equivalent of the "i" leader that was described for Ad2 early L1 transcripts (Alestrom *et al.,* 1982).

OAV promoter regions

The locations of the promoters described in this work are approximate and should be confirmed by more detailed studies of individual transcription units. Nevertheless, inspection of the sequences near the transcription start sites reveals features consistent with the presence of promoters. Possible TATA box (TTAAAA) and CAT boxes (CAAT) are located 30–40 and ~100 nucleotides, respectively, upstream of LHP2.2 transcription start sites. The sequences tGGAAtTtACA and AGGAAGgGcCA are present in the OAV genome at nucleotides 64 and 277, respectively. These motifs are similar to the duplicated enhancer sequence described for the Ad2 E1A region (Hearing and Shenk, 1983) and could influence transcription from the LHP2 promoter that may lie in the region of nucleotides 455–515. The location of the LHP1 promoter has not been determined, but it could conceivably lie within the 46-bp inverted terminal repeat sequence, where an ATATATAA sequence and a possible SP1 transcription factor binding site, GGGCGtG, are present. Similarly, at the RHE the promoter may also incorporate the ITR sequence because there are no other candidate TATA box sequences between the ITR and the first Met, which is located only 117 nucleotides from the end of the genome. Alternatively, these promoters might lack TATA box sequences. For the P32 promoter a TATA box is found ~80 nucleotides upstream of the P32.1 transcription start site and a CAT box is present ~60 nucleotides

further upstream. The Ad2 IVa_2 promoter has an initiator sequence and a downstream TATA element (Carcamo *et al.,* 1991; Chen *et al.,* 1994). For the OAV IVa_2 promoter region there are possible TATA elements 90 and ~140 nucleotides upstream of the start sites. Other features in the proposed promoter region are not readily identifiable and other studies will be required to locate it more precisely. The E2 promoter region contains a TATA box ~25 nucleotides from the transcription start site and may contain binding sites for ATF and SP1 transcription factors but apparently lacks a binding site for E2F, a transcription factor that is important in controlling expression of this region during Ad2 infection. In the E4 region, classic TATA box motifs are present 50 nucleotides upstream of the E4P1 start site and very close to the start sites for E4P2 transcripts. For the latter, the Ad2 E1A enhancer motif AGGAAGTtAaA is also present nearby. This is consistent with the apparent greater strength of the E4P2 promoter. These observations encourage more detailed studies to define promoter regions by subcloning/reporter gene expression and footprint analyses. More detailed temporal studies are also need to define the promoter(s) that is activated immediately after infection.

Comparison with EDS

In the proposed third virus group, EDS is the only other member for which the complete nucleotide sequence has been determined (Hess *et al.,* 1997). The elucidation of the OAV transcription map highlights features of the two genomes that merit brief comment. At the LHE, EDS has no reading frame that matches the OAV 9.7-kDa ORF, although other ORFs, including a P32K homologue are present. In the E2 region, the Met adjacent to the protease cleavage site in pTP is not conserved, although the cleavage site is present. An alternative Met is present 14 residues upstream. Whether alternative splicing occurs during EDS infection remains to be determined. In the E4 region, both viruses contain two ORFs which have homology with each other (our unpublished results) and with the Ad2 E4 34-kDa protein (Vrati *et al.,* 1996b). However, OAV has an additional 171-kDa ORF that is missing in EDS. In both viruses there is a region of ~1 kb between E4 and ORFs to the right. The prediction is that this region of EDS may also contain a promoter(s), as demonstrated for OAV. At the RHE OAV contains six ORFs, of which four are closely related but of unknown function (Xu *et al.,* 1997). EDS contains two ORFs that have homology with each other and to the four ORFs of OAV. As the OAV genome (29.58 kb) is shorter than EDS (33.7 kb), the VA RNA gene that lies to the right among other EDS-specific ORFs is missing from OAV (Venktesh *et al.,* 1997).

Promoter function in other cells

Several of the transcripts defined in this study were used to monitor OAV promoter function in OAV-infected human cells (Khatri *et al.,* 1997). A range of promoter function was observed, depending on the cell type. In MRC-5 lung fibroblasts, for example, all known promoters were active (IVa_2 was not examined). However, in HepG2 liver carcinoma cells, none of the promoters was active. The MLP was essentially inactive in all cell types, probably because DNA replication did not occur (Khatri *et al.,* 1997). The characterisation of OAV promoters and transcripts will facilitate additional studies in the assessment of recombinant OAV as a potential gene therapy vector.

MATERIALS AND METHODS

Virus and cell culture

OAV287 (OAV) (Peet *et al.,* 1983; Boyle *et al.,* 1994) was propagated in CSL503 cells, a sheep foetal lung cell line, as described (Boyle *et al.,* 1994). CSL503 cells were maintained in EMEM plus 10% FCS.

mRNA preparation and PCR

mRNA was prepared from OAV-infected (M.O.I. of 20) or mock-infected CSL503 cells at various times between 12 and 48 h postinfection (pi) using Quickprep micro-mRNA synthesis kits incorporating selection by oligo(dT) (AMRAD/Pharmacia Biotech, Melbourne, Vic). mRNA samples were treated with RNase-free DNase I (Boerhinger Mannheim, North Ryde, NSW) before use for cDNA synthesis. cDNA was prepared using oligo$(dT)_{15}$as a primer for AMV reverse transcriptase (1st Strand cDNA Synthesis Kit, Boehringer Mannheim). In some cases priming with random hexamers was used. Oligonucleotide primers for PCR were designed using the known OAV sequence (Vrati *et al.,* 1995, 1996a, b) and produced using a DNA synthesiser (Applied Biosystems, Foster City, CA). PCR was performed under commonly used conditions (Sambrook *et al.,* 1989) for 40 cycles using annealing temperatures that were close to the melting temperatures of particular primer pairs. The integrity and yield of mRNA and cDNA preparations was monitored by PCR amplification of glyceraldehyde-phosphate dehydrogenase cDNA as described previously (Khatri *et al.,* 1997).

RLM–RACE (RNA ligation-mediated rapid amplification of cDNA ends) (Schaeffer, 1995) together with RT–PCR (collectively referred to here as RACE–PCR) was used to map the 5′ ends of various viral mRNAs. Briefly, mRNAs were digested with alkaline phosphatase (Boehringer-Mannheim) and intact mRNAs were decapped with tobacco acid pyrophosphatase (Epicentre Technologies, Madison, WI) leaving only the full-length mRNAs with a 5′ phosphate group. A single-stranded DNA oligonucleo-

tide with an RNA base at its 3′ end (sequence d(AGC TTA AGC CTC TGA AGG TTAA) rA) (the anchor oligo) was ligated to the 5′ end of mRNAs using T4 RNA ligase (Boehringer-Mannheim). cDNA was prepared from 24-h-pi-polyadenylated mRNA using random primers (Promega, Madison, WI). Subsequent PCR amplification of virus-specific transcripts was carried out using the 5′ anchor oligo in conjunction with a 3′ gene-specific primer. In some cases a nested product was amplified by using an internal 3′ oligonucleotide. RNA molecules that did not accept the anchor oligo, or cDNA molecules that were not fully transcribed could not be amplified by RACE–PCR. Fragments obtained by PCR or RACE–PCR were blunt end-cloned into *Sma*I-cut, phosphatased plasmid pGEM7Zf(−) (Promega, Madison, WI). Clones of the expected size were identified by restriction endonuclease digestion and sequenced using an ABI Prism Model 377 sequencer.

ACKNOWLEDGMENTS

The authors thank A. Mir for providing the anchor oligonucleotide and Drs. P. Molloy and Z. Xu for comments on the manuscript. A.K. acknowledges the support of CSIRO during these studies for her Ph.D.

REFERENCES

Alestrom, P., Akusjarvi, G., Pettersson, M., and Pettersson, U. (1982). DNA sequence analysis of the region encoding the terminal protein and the hypothetical N-gene product of adenovirus type 2. *J. Biol. Chem.* 257, 13492–13498.

Baker, C., and Ziff, E. (1981). Promoters and heterogeneous 5′ termini of the messenger RNAs of adenovirus serotype 2. *J. Mol. Biol.* 149, 189–221.

Boyle, D. B., Pye, A. D., Kockerhans, R., Adair, B. M., Vrati, S., and Both, G. W. (1994). Characterisation of Australian ovine adenovirus isolates. *Vet. Microbiol.* 41, 281–291.

Carcamo, J., Buckbinder, L., and Reinberg, D. (1991). The initiator directs the assembly of a transcription factor IID-dependent transcription complex. *Proc. Natl. Acad. Sci. USA* 88, 8052–8056.

Chen, H. J., Vinnakota, R., and Flint, S. J. (1994). Intragenic activating and repressing elements control transcription from the adenovirus IVA(2) initiator. *Mol. Cell Biol.* 14, 676–685.

Chiocca, S., Kurzbauer, R., Schaffner, G., Baker, A., Mautner, V., and Cotten, M. (1996). The complete DNA sequence and genomic organization of the avian adenovirus CELO. *J. Virol.* 70, 2939–2949.

Harrach, B., Meehan, B. M., Benko, M., Adair, B. M., and Todd, D. (1997). Close phylogenetic relationship between egg drop syndrome virus, bovine adenovirus serotype 7, and ovine adenovirus strain 287. *Virology* 229, 302–306.

Hearing, P., and Shenk, T. (1983). The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell* 33, 695–703.

Hess, M., Bloecker, H., and Brandt, P. (1997). The complete nucleotide sequence of the egg drop syndrome virus: An intermediate between mastadenoviruses and aviadenoviruses. *Virology* 238, 145–156.

Horwitz, M. S. (1990). Adenoviridae and their replication. *In* "Virology" (B. N. Fields and D. M. Knipe, Eds.), pp. 1679–1721. Raven Press, New York.

Kato, S., Sekine, S., Oh, S., Kim, N., Umezawa, Y., Abe, N., Yokoyama-Kobayashi, M., and Aoki, T. (1994). Construction of a human full-length cDNA bank. *Gene* 150, 243–250.

Khatri, A., Xu, Z. Z., and Both, G. W. (1997). Gene expression by atypical recombinant ovine adenovirus vectors during abortive infection of human and animal cells in vitro. *Virology* 239, 226–237.

Kreivi, J., and Lamond, A. (1996). RNA splicing: Unexpected spliceosome diversity. *Curr. Biol.* 6, 802–805.

Maruyama, K., and Sugano, S. (1994). Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 138, 171–174.

Maxfield, L. F., and Spector, D. J. (1997). Readthrough activation of early adenovirus E1b gene transcription. *J. Virol.* 71, 8321–8329.

Natarajan, V., Madden, M. J., and Salzman, N. P. (1984). Proximal and distal domains that control in vitro transcription of the adenovirus IVa2 gene. *Proc. Natl. Acad. Sci. USA* 81, 6290–6294.

Natarajan, V., Madden, M. J., and Slazman, N. P. (1985). Positive and negative control sequences within the distal domain of the adenovirus IVa2 promoter overlap with the major late promoter. *J. Virol.* 55, 10–15.

Norrby, E., Bartha, A., and Boulanger, P. (1976). Adenoviridae. *Intervirol* 7, 117–125.

Osborne, T., Gaynor, R., and Berk, A. (1982). The TATA homology and the mRNA 5′ untranslated sequence are not required for expression of essential adenovirus E1A functions. *Cell* 29, 139–148.

Peet, R. L., Coackley, W., Purcell, D. A., Robartson, C. W., and Micke, B. M. (1983). Adenovirus inclusions in sheep liver. *Aust. Vet. J.* 60, 307–308.

Roberts, R. J., O'Neill, K. E. O., and Yen, C. T. (1984). DNA sequences from the adenovirus 2 genome. *J. Biol. Chem.* 259, 13968–13975.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schaeffer, B. C. (1995). Revolutions in rapid amplification of cDNA ends: New strategies for polymerase chain reaction cloning of full-length cDNA ends. *Anal. Biochem.* 227, 255–273.

Shenk, T. (1996). Adenoviridae: The viruses and their replication. *In* "Fields' Virology" (B. N. Fields, D. M. Knipe, and P. M. Howley, Eds.), pp. 2111–2148. Lippincott–Raven, Philadelphia.

Venktesh, A., Watt, F., Xu, Z. Z., and Both, G. W. (1997). Ovine adenovirus (OAV287) lacks a VA RNA gene. *J. Gen. Virol.* 79, 509–516.

Vrati, S., Boyle, D., Kocherhans, R., and Both, G. W. (1995). Sequence of ovine adenovirus homologs for 100K hexon assembly, 33K, pVIII, and fiber genes: Early region E3 is not in the expected location. *Virology* 209, 400–408.

Vrati, S., Brookes, D. E., Boyle, D. B., and Both, G. W. (1996a). Nucleotide sequence of ovine adenovirus tripartite leader sequence and homologues of the IVa2, DNA polymerase and terminal proteins. *Gene* 177, 35–41.

Vrati, S., Brookes, D. E., Strike, P., Khatri, A., Boyle, D. B., and Both, G. W. (1996b). Unique genome arrangement of an ovine adenovirus: Identification of new proteins and proteinase cleavage sites. *Virology* 220, 186–199.

Webster, A., Leith, I. R., and Hay, R. T. (1994). Activation of adenovirus-coded protease and processing of preterminal protein. *J. Virol.* 68, 7292–7300.

Xu, Z. Z., Hyatt, A., Boyle, D. B., and Both, G. W. (1997). Construction of ovine adenovirus recombinants by gene insertion or deletion of related terminal region sequences. *Virology* 230, 62–71.